

Published in final edited form as:

*Psychol Assess*. 2016 May ; 28(5): 539–548. doi:10.1037/pas0000139.

# Reliability, Factor Structure, and Measurement Invariance of the Dominic Interactive Across European Countries: Cross-Country Utility of a Child Mental Health Self-Report

**Rowella C. W. M. Kuijpers**,
Radboud University Nijmegen

**Roy Otten**,
Radboud University Nijmegen

**Ad A. Vermulst**,
Radboud University Nijmegen

**Adina Bitfoi**,
The Romanian League for Mental Health, Bucharest, Romania

**Dietmar Goelitz**,
Leuphana University of Lueneburg

**Ceren Koç**,
Yeniden Health and Education Society, Istanbul, Turkey

**Zlatka Mihova**,
New Bulgarian University

**Ondine Pez**,
Paris Descartes University

**Mauro Carta**,
University of Cagliari

**Katherine Keyes**,
Columbia University

**Sigita Lesinskiene**,
University of Vilnius

**Rutger C. M. E. Engels**, and
Radboud University Nijmegen

**Viviane Kovess**
Paris Descartes University

Correspondence concerning this article should be addressed to Rowella C. W. M. Kuijpers, Behavioural Science Institute, Radboud University Nijmegen, P.O. Box 9104, 6500 HE Nijmegen, the Netherlands. r.kuijpers@pwo.ru.nl.

The authors have declared that there are no conflicts of interest.

Supplemental materials: http://dx.doi.org/10.1037/pas0000139.supp

## Abstract

Large-scale international surveys are important to globally evaluate, monitor, and promote children's mental health. However, use of young children's self-reports in these studies is still controversial. The Dominic Interactive, a computerized *DSM-IV*–based child mental health self-report questionnaire, has unique characteristics that may make it preeminently appropriate for usage in cross-country comparisons. This study aimed to determine scale score reliabilities (omega) of the Dominic Interactive in a sample of 8,135 primary school children, ages 6–11 years old, in 7 European countries, to confirm the proposed 7-scale factor structure, and to test for measurement invariance of scale and item scores across countries. Omega reliability values for scale scores were good to high in every country, and the factor structure was confirmed for all countries. A thorough examination of measurement invariance provided evidence for cross-country test score comparability of 5 of the 7 scales and partial scale score invariance of 2 anxiety scales. Possible explanations for this partial invariance include cross-country differences in conceptualizing items and defining what is socially and culturally acceptable anxiety. The convincing evidence for validity of score interpretation makes the Dominic Interactive an indispensable tool for cross-country screening purposes.

### Keywords

child self-report; the Dominic Interactive; construct validity; measurement invariance; cross-country research

---

Interest in globally monitoring the mental health of children to prevent the occurrence and exacerbation of child mental health problems is increasing. The National Research Council and Institute of Medicine (O'Connell, Boat, & Warner, 2009) presented prevalence estimates of the most common mental health problems in children and adolescents from over 50 community studies across the world. These numbers revealed that 17% of the children and adolescents had one or more emotional or behavioral disorder (anxiety 8.0%, depression 5.2%, disruptive disorders 6.1%, attention-deficit/hyperactivity disorder [ADHD] 4.5%, substance use disorder 10.3%). Also, comorbidity of these disorders appeared to be high, varying from 3–11%. Several studies have shown that the first symptoms of these mental health problems have their onset in childhood and that the prevalence of these subthreshold levels and comorbid manifestations is even higher, with estimates above 20% (Egger & Angold, 2006; Merikangas, Nakamura, & Kessler, 2009). In addition, if left untreated, there is a high risk of ongoing mental health problems or the occurrence of other mental health problems throughout adolescence and adulthood (Costello, Egger, & Angold, 2005; Costello, Mustillo, Erkanli, Keeler, & Angold, 2003; Ford, 2008; Reef, Diamantopoulou, van Meurs, Verhulst, & van der Ende, 2009). Furthermore, children with mental health problems in need of care are often not recognized or do not receive the necessary care (Jensen et al., 2011; Sourander et al., 2005). These findings warrant monitoring and evaluating child mental health and early screening of emerging mental health problems on an international level.

The World Health Organization, UNICEF, and the European Union spend considerable funding on large-scale health surveys to gain insight into the well-being of young children

across countries. These surveys are designed to obtain information on the prevalence and developmental course of mental health problems, related risk factors, and available care (Rutkowski & Svetina, 2014). Because a multimethod and multi-informant approach is preferred, these studies have used various sources of information, such as national statistics (e.g., number of children seeking treatment, money spent on child mental health care, child abuse, police reports) and information from parents, teachers, or clinicians. Although most scholars agree about the importance of using children to provide information on their own mental health status, there are still reservations about the reliability of such information due to limited language skills, cognitive skills, and self-reflective abilities (Fallon & Schwab-Stone, 1994). Despite the increase in available self-report instruments, specifically for emotional problems (Silverman & Ollendick, 2005), no instruments have been developed to screen for a wide array of mental health problems (i.e., emotional *and* behavioral problems) in primary school children.

The Dominic Interactive, a pictorial computerized self-report instrument for screening mental health problems in primary school children (Valla, 2000), has been shown to be a valid alternative that partly overcomes these barriers. Traditionally, there are two approaches to conceptualize child mental health problems. The *categorical approach* is based on expert consensus about disorders containing distinctive combinations of symptoms. According to the *dimensional approach*, syndromes are derived from empirically co-occurring symptoms that move along a continuum from normal to clinical (Hartman et al., 2001; Lahey et al., 2004). The Dominic Interactive is based on the categorical approach as reflected by the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-IV*; American Psychiatric Association, 1994; *DSM-5*; American Psychiatric Association, 2013) and measures the symptoms of the seven most prevalent mental health problems in primary school children, including both emotional *and* behavioral syndromes (i.e., specific phobia [SPh], separation anxiety disorder [SAD], generalized anxiety disorder [GAD], major depressive disorder [MDD], oppositional defiant disorder [ODD], conduct disorder [CD], and ADHD). The items of the Dominic Interactive cover each of the symptoms as described in *DSM*, except for two that were considered to be developmentally inappropriate. All items (both text and pictures) went through an extensive selection procedure entailing clinical expert judgment, qualitative child evaluations, and statistical testing (Valla, 2000; Valla, Bergeron, Bérubé, Gaudet, & St-Georges, 1994; Valla, Bergeron, Bidaut-Russell, St-Georges, & Gaudet, 1997). Ten items that reflect positive behaviors comprise the Strengths and Competencies scale. Because these items are merely added to increase acceptability, they are not considered part of the original conceptual framework. By omitting time-related criteria (e.g., onset, duration, frequency), severity, and interference, and by displaying the symptoms using pictures of daily situations combined with both written and spoken sentences, some of the previously mentioned age-related barriers are overcome. This approach, however, also limits the utility of the Dominic Interactive to the sole purpose of screening for potential mental health problems and it is not meant to diagnose disorders. The Dominic Interactive, therefore, uses “tendencies towards” when referring to scale scores.

Several properties of the Dominic Interactive may even increase its utility in international survey studies. First, the instrument combines different modalities (i.e., text, pictures, and voice-over) to display symptoms, decreasing the possible risk of translation errors and

subsequent cultural differences in the interpretation of the item content. Second, the ethnicity of the character can be adapted to White, Hispanic, Asian, or African, and each character has a corresponding ethnic and gender-neutral name to optimize identification of the child. Third, the Dominic Interactive is currently available in 11 different languages (e.g., English, French, Spanish, Turkish) with voice-overs from native speakers.

Although these characteristics of the Dominic Interactive may facilitate its usage in international studies, other challenges include cultural differences in understanding the underlying mental health constructs and interpreting the items that reflect the construct of interest (Rutkowski & Svetina, 2014; van Widenfelt, Treffers, de Beurs, Siebelink, & Koudijs, 2005). These challenges affect the comparability of scale scores across groups, making it imperative that latent variables are well operationalized and measured equivalently across countries, which is referred to as *measurement invariance* (Sass, 2011).

Without convincing evidence that measures have the same meaning across groups, valid conclusions about differences between groups are unwarranted. The scale scores of an instrument should meet the requirements of measurement invariance on at least three levels (Chen, 2007; Rutkowski & Svetina, 2014; Sass, 2011; Schmitt & Kuljanin, 2008). First, *configural invariance* requires the same items to be related to the same factors, indicating that each construct reflects the same underlying concept. Second, *metric invariance* refers to the equality of factor loadings of the items belonging to the construct, implying that all items contribute equally strongly to the conceptual meaning of the scale. Third, *scalar invariance* refers to the equivalence of thresholds of the items across groups, which makes it possible to compare mean differences. When measurement invariance is supported at a configural, metric, and scalar level, differences in scale scores can be attributed to authentic differences between groups on the latent variable of interest, and not to cross-country measurement differences, translation errors, or biases. The majority of studies on measurement invariance have involved subjects concerning human resources and organizational research (Schmitt & Kuljanin, 2008). In international child mental health studies, this psychometric concept has received little attention, and the available research has often been limited to a few groups and small sample sizes (Chen, 2007). To our knowledge, the only available large cross-country study on measurement invariance of child mental health constructs used parent report (Child Behavior Checklist) and tested for configural invariance, but not metric and scalar invariance (Ivanova et al., 2007). It is therefore imperative to establish the construct validity and measurement invariance of scale and item scores of the Dominic Interactive to be able to use this screening tool for cross-country comparison.

Although many psychometric properties of the Dominic Interactive (e.g., internal consistency, test–retest reliability, criterion validity) have been examined with satisfactory-to-good results in various countries, such as Canada, the United States, France, and the Netherlands (Bergeron, Berthiaume, St-Georges, Piché, & Smolla, 2013; Kuijpers, Otten, Krol, Vermulst, & Engels, 2013; Kuijpers, Otten, Vermulst, & Engels, 2014; Scott, Short, Singer, Russ, & Minnes, 2006; Shojaei et al., 2009; Valla et al., 2002), there is a paucity of studies on the construct validity and measurement invariance of item and scale scores. Available studies were conducted in only two countries (Canada and the Netherlands) and only in small samples, hampering a proper item-level confirmatory analysis (CFA). Instead,

these studies performed exploratory factor analysis at a scale level, both revealing a two-factor structure, internalizing and externalizing problems (Dugré, Trudel, & Valla, 2001; Kuijpers et al., 2013). The latter research group also conducted a one-factor CFA for each subscale, finding good results for five of seven scales. The only proper CFA has been performed by Kuijpers et al. (2014) on the Dutch sample of the current study. In that study (Kuijpers et al., 2014), we verified the original seven-scale factor structure and confirmed measurement invariance of scale scores for sex and age, but we did not test for cross-country measurement invariance. Conclusively, to employ the Dominic Interactive in international epidemiological and prevalence studies, the instrument would greatly benefit from examining the factor structure in various countries and from determining measurement invariance of scale constructs across countries.

In sum, the purpose of the present study was to compute scale score reliabilities, test the proposed seven-scale factor structure, and, subsequently, to focus extensively on testing measurement invariance of scale and item scores of the Dominic Interactive in primary schoolchildren (6–11 year olds) across seven European countries. We expected to find good reliabilities for all scale scores in each country, with omega values exceeding alpha values (cf. Kuijpers et al, 2014). Furthermore, we expected to find evidence for the seven-scale factor structure in each country, although, considering the high level of comorbidity that has been established in prior prevalence and epidemiological studies (Egger & Angold, 2006; Merikangas et al., 2009), we did not presume the constructs to be independent. Finally, we did not have a clear expectation about measurement invariance, because no other large-scale study of a child self-report screening instrument has been conducted before.

## Method

### Sample and Participants

The School Children Mental Health in Europe (SCMHE) project was a cross-sectional European study ($N$ = 8,135) conducted in seven countries: Bulgaria ($n$ = 1,385), Germany ($n$ = 894), Italy ($n$ = 757), Lithuania ($n$ = 1,278), Netherlands ($n$ = 1,503), Romania ($n$ = 1,397), and Turkey ($n$ = 921). The purpose of this project was to establish a set of indicators to monitor children's mental health, to determine major risk factors, and to screen children for the most common mental health problems. The European Commission funded the project. In each country or in the participating regions of a country (e.g., Sardinia in Italy), primary schools were randomly selected. In each school, classes were randomly selected and in each class there was a random selection of five to six. The aim was to select 48 children in each school, 10 in each grade (five per class) for schools with five grades and 12 in each grade (six per class) for schools with four grades. In the Netherlands and Germany, more children per school were asked to participate, due to lower school participation. Mean age was 8.7 years ($SD$ = 1.44), with 97.7% falling in the 6–11 year age range. Boys and girls were represented equally (51.3% boys and 48.7% girls). The vast majority of children had a mother or father whose nationality was the country of residence, 97.2% and 96.5%, respectively. Statistics on German nationality were not available. In 53.8% of the families, at least one parent completed college or technical school, 35.7% completed secondary school, and only a minority (10.5%) had some primary or secondary education. Regarding

community size, 34.2% of the children came from a rural area or small village, 24.6% lived in a small- or middle-sized town, and 41.2% lived in a large town. Community size and educational level were not distributed equally across countries. Descriptive statistics are presented in Table 1. The parents provided information on community size, nationality, and educational level. Due to low parent participation rates compared to child participation rates across the sample, demographics are only available for a sub-sample of children and therefore representativeness cannot be adequately determined. Additional information about the sampling method and the sample composition was described in the final SCMHE report (Kovess et al., 2015). Research in all countries received ethical approval from appropriate authorities.

## Procedure

From the 14,317 children present in the randomly selected schools, parents of 9,084 children gave consent to participate and 8,135 children completed the Dominic Interactive and were included in this study. Main reasons for not being in this study were incomplete data or absence (e.g., illness) on the day of administration. In all countries, parents were informed about the purpose of the study and they could refuse participation by returning the consent letter. Interviewers visited the schools and, in small groups, selected children received a general instruction explaining the purpose of the study (i.e., to find out how primary school children think, act, and feel). They were reassured that their answers were confidential and were not evaluated as right or wrong. In the next step, they were prompted to put on their headphones and start the computer program (the Dominic Interactive). Children were allowed to ask questions, but school investigators were instructed not to suggest potential answers and to give children a sense of privacy by keeping a distance.

First, the child was guided through an example that introduced Dominic and explained the response method, that is, clicking the “Yes” or “No” button by using the mouse. Next, the child was presented with pictures of Dominic engaging in daily situations at home, in school, or with friends, and displaying a symptom of one of the mental health problems. The specific symptom was also printed in text at the bottom of the screen and read aloud by a voice-over, asking the child whether he or she ever felt, thought, or acted like Dominic. The administration took 10–20 min, depending on the child's age.

## Measures

The Dominic Interactive is a child self-report screening instrument comprising 81 items that represent symptoms of the seven most prevalent *DSM-IV* mental health problems in school children and also match the *DSM-5* criteria for: SPh, SAD, GAD, MDD, ODD, CD, and ADHD. The Strengths and Competencies scale consists of 10 positive items (Valla, 2000). The instrument is fully computerized, enabling children to independently answer the questions, automatically saving the answers (“Yes” coded as 1 and “No” as 0) and summing up the scale scores. Comparison with *DSM-5* showed no differences in item content. See Figure 1 for an example of an item of the Dominic Interactive (Valla, 2000).

## Statistical Analyses

**Reliability—**In the first step, we computed Cronbach's alphas for each of the seven scale scores separately for each country. However, recently, there has been a debate about the usefulness of alpha measure (McDonald, 1999; Sijtsma, 2009). One reason is the required assumption that all items in a scale have the same true score variance (also known as tau-equivalence). Relaxing this assumption in the context of structural equation modeling will give a better estimate for the reliability of scale scores (Bacon, Sauer, & Young, 1995). We computed McDonald's omega as an alternative reliability measure (Kuijpers et al., 2014; Stone et al., 2013). Consistent with these studies, omega ($\omega_h$) was computed from one-factor solutions by conducting CFA for each scale.

**Factor structure—**Because the theoretical framework of the Dominic Interactive is based on the seven most common *DSM-IV* mental health problems, we performed a categorical confirmatory factor analysis (CCFA) on the item scores of these seven scales and excluded the Strength and Competencies scale because it was not part of the proposed theoretical framework. With CCFA, the binary response scales of the items are represented by thresholds; the input for the analysis is a matrix with tetrachoric correlations and thresholds. In accordance with *DSM-IV* syndrome composition (Valla, 2000), some items are assigned to more than one scale (e.g., “irritability” is a symptom of GAD, MDD, and ODD). However, this hinders a simple CFA; therefore, we allocated overlapping items to the original scale. See Appendix A in the supplemental materials for a path model of the hypothesized factor structure. We conducted CFA in Mplus, Version 6.1 (Muthén & Muthén, 1998–2007) using the WLSMV-estimator (i.e., weighted least square estimator with a mean- and variance-adjusted chi-square test statistic) due to binary response categories. We used chi-square values, *p* values, the comparative fit index (CFI), and root mean square error of approximation (RMSEA) to evaluate the goodness of model fit. According to the conventional criteria, a CFI greater than .90 and RMSEA less than .08 are considered acceptable and a CFI greater than .95 and RMSEA less than .05 are considered good (Marsh, Hau, & Wen, 2004). Because children in each country are nested within classes and classes within schools, the possibility exists that the data are nonindependent. To correct for nonindependence, we applied the COMPLEX procedure in Mplus to get unbiased estimates of the standard errors. Only the standard errors are affected, and they become slightly higher compared with the uncorrected results. Due to software limitations, it was not possible to correct for schools and classes simultaneously. We compared the results corrected for schools with the results corrected for classes and found no differences in significance levels of the parameter estimates. The results corrected for schools are presented in this article.

First, we tested the seven-factor solution for each country separately. Next, to test measurement invariance of scale constructs, we followed a three-level process to explore in depth the proposed underlying *DSM-IV* structure across countries. First, we tested configural, metric, and scalar invariance across countries for the *overall* seven-scale structure, following Steenkamp and Baumgartner's (1998) recommendations. Configural invariance measures overall model fit of the underlying latent factors, assuming an equal pattern of zero and nonzero loadings across countries. Metric and scalar invariance were tested simultaneously because of the binary (categorical) nature of the variables (Muthén &

Muthén, 1998–2007; Kim & Yoon, 2011). Metric invariance implies equal factor loadings across countries. Scalar invariance implies equal thresholds across countries. With a chi-square difference test (DIFFTEST in Mplus), the fit of the configural model was compared with the (equal loading and threshold) constrained model. A significant chi-square indicates inequality of factor loadings and thresholds. However, scale length, sample size, and number of groups have a direct effect on the chi-square difference test and, as a result, trivial differences may be specified as a model misfit (Chen, 2007; Rutkowski & Svetina, 2014). Therefore, we conformed to the recommendations of Chen (2007) and Cheung and Rensvold (2002) that invariance should not be rejected if the decrease in CFI is less than .01 and the increase in RMSEA is less than .015.

Global fit measures can mask the invariance of individual items (e.g., if all items of a scale, except for one, show perfect measurement invariance, global fit measures will still show acceptable values; Schmitt & Kuljanin, 2008). Therefore, we repeated the measurement invariance tests across countries for each scale separately (measurement invariance per scale). To get a more detailed picture to detect items responsible for noninvariance, we followed a two-step procedure. First, we computed a *fully* constrained model (implying equality of factor loadings and thresholds of *all* items for each scale). Next, we relaxed factor parameters for each item separately (i.e., allowed the factor loading and threshold of that specific item to vary freely) and compared the relaxed model with the *fully* constrained model. CFI changes exceeding the .01 criterion and RMSEA changes exceeding the .015 criterion were indicative of noninvariance of an item across countries (measurement invariance per item).[1]

## Results

### Reliability

Alpha coefficients for all scale scores were consistent across countries, showing moderate-to-acceptable values for most scale scores (see Table 2). Good values (>.80) were primarily found for scales scores for MDD, CD, and ADHD. Lowest alpha coefficients were found for scale scores for SPh and SAD (range: .54–.75). Omega values for all scale scores appeared to be in a good-to-high range (range: .80–.97) across countries. Only six values were below .80, but still in an acceptable range (range: .75–.79).

### Construct Validity: Factor Structure for Each Country

For each country, the seven-factor model showed acceptable fit measures, with CFI values ranging from .894–.946 and RMSEA values ranging from .011–.016, confirming the proposed underlying factor structure. Table 3 shows Pearson correlations among the seven scales for all seven countries and the range of correlations across countries. As expected, results showed substantial correlations between all scales with the exception of CD with the anxiety scales (SPh, SAD, and GAD). MDD also had substantial correlations with the anxiety scales and with ODD, CD, and ADHD.

[1]The syntaxes of all analyses can be requested from the corresponding author.

### Construct Validity: Overall Measurement Invariance

The baseline model (factor loadings and thresholds free to vary across the seven countries) showed an acceptable-to-good model fit, $\chi^2(21966) = 26{,}305.75$, $p = .000$, CFI = .928, RMSEA = .014. This indicates that configural invariance of constructs across countries is supported by acceptable CFI values and good RMSEA values, confirming the similarity of the seven-scale factor structure across countries. The constrained model also showed an acceptable-to-good model fit, $\chi^2(22410) = 27{,}378.45$, $p = .000$, CFI = .918, RMSEA = .014. The chi-square difference test showed that the constrained model differed significantly from the baseline (configural) model, $\Delta\chi^2(444) = 2{,}656.98$, $p = .000$, although changes in CFI (ΔCFI = −.010) and RMSEA (ΔRMSEA = .001) were within the acceptable limits of .01 and .015, respectively. This indicates that metric invariance and scalar invariance were supported, confirming the equivalence of factor loadings and thresholds of items across countries.

### Construct Validity: Measurement Invariance per Scale

We also tested measurement invariance for each scale separately across countries. Table 4 shows that all fit indices for configural invariance were acceptable to good, with CFI values ranging from .932–.990 and RMSEA values ranging from .017–.053. Constraining factor loadings and thresholds to be equal across countries yielded acceptable-to-good CFI values and good RMSEA values for five of the seven constructs. The SAD and GAD scales showed unacceptable fit values.

Although the fit measures of the constrained model were satisfactory for five of the seven constructs, the *changes* in these fit indices compared with the baseline model (ΔCFI and ΔRMSEA) showed ambiguous results. The changes in CFI values of six of seven scales exceeded the recommended criterion of .01 for every scale except for CD, implying that the scale scores of these six scales were noninvariant. However, the changes in RMSEA for five of seven scales were below .015, except for SAD and GAD, indicating that only these two constructs were noninvariant. Applying the most stringent criteria, that is, invariance meeting the recommended standard of ΔCFI *as well as* ΔRMSEA, all constructs except CD would be noninvariant across countries.

### Construct Validity: Measurement Invariance per Item

To obtain a more detailed picture of the ambiguity of the measurement invariance results for each scale, we analyzed which item scores contributed significantly to changes in CFI and RMSEA by comparing the fully constrained model of each scale (i.e., equal loadings and thresholds across countries) with a relaxed model (i.e., factor loading and threshold of one item at the time to vary freely). Examining the changes in CFI, nine items (three belonging to SAD and six belonging to GAD) exceeded the standard of .01, indicating that these item scores significantly contributed to the metric and scalar noninvariance of that scale construct. None of the ΔRMSEA values exceeded the standard value of .015. Changes in CFI and RMSEA for all items of the other scales were within the acceptable limits of .01 and .015, respectively, suggesting that these item scores are invariant at a metric and a scalar level. Table 5 displays the results of the chi-square difference tests for each item of SAD and

GAD, including the factor loadings and thresholds of the baseline model. See Appendix B in the supplemental materials for the results of items of the other scales.

It was not possible to test whether factor loadings or thresholds caused the noninvariance, because for this type of (binary) data, both parameters must be estimated simultaneously. However, computing the variance of the factor loadings and the thresholds for each item of SAD and GAD across the seven countries showed that the factor loadings varied from .00–.04 ($M$ = .01) and that the thresholds varied from .00–.30 ($M$ = .07). This suggests it may be more plausible that differences in thresholds rather than differences in factor loadings caused noninvariance of these items.

Based on the integration of the results from the three-level process while applying the most stringent criteria (i.e., both ΔCFI *and* ΔRMSEA may not exceed the recommended cutoffs for items and scales to be defined as invariant), we concluded that five of seven scale constructs were invariant across countries, and SAD and GAD were partial invariant (i.e., some items are invariant but some are not).

## Discussion

This study aimed to test the scale score reliability and to confirm the proposed *DSM-IV*–based seven-scale factor structure of the Dominic Interactive, a child self-report screening instrument, across seven countries. Next, we conducted detailed analyses to examine measurement invariance of the seven scale constructs. The scale score reliability of all seven scales of the Dominic Interactive appeared to be good in each country, with omega values exceeding alpha values, which indicates that relaxing the assumption of identical true score variance of items will provide better reliability estimates, as suggested by other scholars (Muthén, 1984; Revelle & Zinbarg, 2009; Stone et al., 2013). The proposed *DSM-IV*–based seven-scale factor structure of the Dominic Interactive (Valla, 2000) was supported in each country, implying that the same conceptual model underlies the Dominic Interactive within each country.

Regarding measurement invariance, the conceptual model was similar across countries for five of the seven scale constructs (i.e., SPh, MDD, ODD, CD, and ADHD), indicating that the items of these scales are related equally strongly (metric invariance) to the same latent construct (configural invariance) and item thresholds are similar across countries (scalar invariance). This justifies the comparison of scores of these five scales across the seven European countries, thereby increasing the prospects of applying the Dominic Interactive in cross-country prevalence and epidemiology studies. The partial construct invariance of SAD and GAD across countries appeared to be related to the variability in thresholds of specific items rather than in factor loadings. This means that the average scale score is higher in one country than another country. Comparing the prevalence rates of separation and generalized anxiety across countries is therefore precarious. One possible explanation concerns the meticulous translation from the original French and English items into other European national languages and to accurately operationalize and reflect these inner mental states (Sass, 2011; van Widenfelt et al., 2005). In particular, the GAD scale contains items that are difficult to operationalize and differentiate from other factors, as was shown in prior studies

(Kuijpers et al., 2013; Kuijpers et al., 2014). A second possible explanation is that the conceptual meaning of separation and generalized anxiety may vary across countries, depending on the cultural context of reference and what is considered socially acceptable or problematic fear (Sass, 2011; van Widenfelt et al., 2005). These explanations are referred to as linguistic and conceptual equivalence, respectively, and are intertwined with statistical types of equivalence as tested in the present study (Ho et al., 2014). Although the guidelines for the development and translation of a test emphasize that translation of the gist of the concepts is more important than literal word-to-word translation, this might also lead to an increased risk of errors and differences in conceptual meaning and interpretation of items (Gudmundsson, 2009; van Widenfelt et al., 2005). Future studies should reflect on these complementary types of equivalence, because limited linguistic and conceptual equivalence might account for the presently found differences between countries in the anxiety scale scores.

The literature provides no specific guidelines on how to deal with partial invariance, that is, what to do when some item scores are invariant and some are not, and how this affects the reliability of the scale scores and the validity of the meaning of the scale, and the decisions that are made based on this (Millsap & Kwok, 2004; Schmitt & Kuljanin, 2008). The proposed options (e.g., deleting items, retaining all items, quit using the entire scale) all depend on other aspects, for example, the amount of noninvariance, the number of noninvariant items, and the purpose of the measure (Millsap & Kwok, 2004; Sass, 2011). Exploring the source of the noninvariance and its impact on the outcomes can be an important next step to substantiate this decision-making process (Schmitt & Kuljanin, 2008). Although noninvariance of some items does not necessarily imply the entire scale is unusable, caution is warranted when making cross-country comparisons, and possible inference problems, biases, or errors should be considered (Millsap & Kwok, 2004; Sass, 2011).

The limitations of this study are predominantly methodological and statistical. First, the national samples may not be completely representative, because, in some countries, the sampling procedure was restricted to certain areas (e.g., the Istanbul area in Turkey), or was adapted due to a limited number of schools willing to participate (e.g., in the Netherlands and Germany). This could have influenced the measurement invariance of scale constructs because associations between items and their underlying latent variable may vary not only across countries, but also across subpopulations. In addition, ongoing immigration of people from outside Europe and the migration and increased blending of European cultures and subpopulations also might have influenced measurement invariance of scale constructs, because differences between countries might be less pronounced. However, this study was a first attempt in determining whether the Dominic Interactive could be useful as a screening tool for mental health problems of children living in various European countries and whether, in general, the scales and scale scores could be comparable at *country* level. Testing whether there are differences in interpretation of scale scores between the dominant *culture* and *subcultures* or between various *ethnicities* requires different selection procedures. The present sample size was too small to create relevant subgroups and was not selected for this purpose. This would be of great interest for future research.

A second limitation is that the European countries in this study do not represent Europe in total. Specific regions were not represented (e.g., Scandinavian countries) or underrepresented (e.g., Southern Europe). This warrants careful attention when generalizing the present results to other European countries. A final statistical limitation does not necessarily concern the present study directly but relates to the general lack of available research on measurement invariance of the proposed constructs as measured by screening instruments in the field of child mental health (Schmitt & Kuljanin, 2008). Currently available research is based mainly on simulation studies or conducted predominantly on a limited number of groups, items, and factors; small sample sizes; continuous item responses; and normally distributed data. In the present study, we had a sample of over 8,000 cases, seven groups, seven factors, 81 items, categorical item responses, and skewed data distributions. This makes it difficult to determine whether the performance of the fit statistics in the present empirical study is comparable to prior findings (Sass, 2011). Consequently, there is a need for more research on measurement invariance of the proposed constructs as measured by screening instruments in international surveys, and we greatly encourage scholars to extend this to the field of child mental health assessment.

Finally, one important theoretical limitation should be mentioned regarding *DSM* as underlying conceptual framework of the Dominic Interactive. Confirming the proposed factor structure in every country does not necessarily imply this is also the *best* or *only* model to conceptualize child mental health problems. There has been discussion whether this framework is applicable to children. Because correlations between specific problem clusters (e.g., anxiety disorders) are generally high, this could imply that mental health problems might be less differentiated in children and it has been argued that a simpler model might more accurately reflect the developmental characteristics of young children (Lahey et al., 2004, 2008). Future studies comparing several theoretically based models (e.g., categorical and dimensional or with a varying number of factors) might give a more detailed insight into the taxonomy of child mental health problems in various countries (cf. Hartman et al., 2001) and might potentially clarify some of the presently found between-country differences in the SAD and GAD scale scores. In addition, the substantial correlations between most scales do not suggest potential second-order factors. However, there also appeared to be differences in interscale correlations between the countries, so exploring comorbidity in child mental health problems on a cross-country level might also be interesting for future studies.

## Conclusion

This study greatly adds to the field of child mental health assessment because it is the first to adopt such a sophisticated procedure to analyze the comparability of a child self-report instrument across countries. This study convincingly showed good results for five of the most common mental health problems in primary school children (i.e., SPh, MDD, ODD, CD, and ADHD) and two scales needing further research (i.e., SAD and GAD). Finally, it is worth mentioning that the item content of the Dominic Interactive scales still matches the criteria of the currently implemented *DSM-5*. Conclusively, this study does not only hold a plea for child mental health research to face the methodological challenges that arise in the

context of doing cross-country surveys, but it also demonstrates the unique properties of the Dominic Interactive and its utility within this specific context.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

Rowella C. W. M. Kuijpers, Roy Otten, and Ad A. Vermulst, Behavioural Science Institute, Radboud University Nijmegen; Ondine Pez, École des Hautes Études en Santé Publique, Paris Descartes University; Adina Bitfoi, The Romanian League for Mental Health, Bucharest, Romania; Mauro Carta, Psychiatric Unit, Department of Public Health, University of Cagliari; Dietmar Goelitz, Center for Applied Sciences of Health, Leuphana University of Lueneburg; Katherine Keyes, Mailman School of Public Health, Columbia University; Ceren Koç, Yeniden Health and Education Society, Istanbul, Turkey; Sigita Lesinskiene, Clinic of Psychiatry, Faculty of Medicine, University of Vilnius; Zlatka Mihova, New Bulgarian University; Rutger C. M. E. Engels, Behavioural Science Institute, Radboud University Nijmegen; and Viviane Kovess, École des Hautes Études en Santé Publique, Paris Descartes University.

The School Children Mental Health in Europe project has been funded by the European Commission (Grant 2006336). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

## References

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th. Washington, DC: Author; 1994.

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th. Washington, DC: Author; 2013.

Bacon DR, Sauer PL, Young M. Composite reliability in structural equation modeling. Educational and Psychological Measurement. 1995; 55:394–406. http://dx.doi.org/10.1177/0013164495055003003.

Bergeron L, Berthiaume C, St-Georges M, Piché G, Smolla N. Reliability, validity, and clinical use of the Dominic Interactive: A DSM-based, self-report screen for school-aged children. Canadian Journal of Psychiatry/Revue Canadienne de Psychiatrie. 2013; 58:466–475. [PubMed: 23972108]

Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. Structural Equation Modeling. 2007; 14:464–504. http://dx.doi.org/10.1080/10705510701301834.

Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling. 2002; 9:233–255. http://dx.doi.org/10.1207/S15328007SEM0902_5.

Costello EJ, Egger H, Angold A. 10-year research update review: The epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden. Journal of the American Academy of Child & Adolescent Psychiatry. 2005; 44:972–986. http://dx.doi.org/10.1097/01.chi.0000172552.41596.6f. [PubMed: 16175102]

Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. Archives of General Psychiatry. 2003; 60:837–844. http://dx.doi.org/10.1001/archpsyc.60.8.837. [PubMed: 12912767]

Dugré S, Trudel M, Valla JP. Individual and cultural considerations concerning the mental health of children: Use of the Dominic-R. Revue Canadienne de Psychoéducation. 2001; 30:119–138.

Egger HL, Angold A. Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. Journal of Child Psychology and Psychiatry. 2006; 47:313–337. http:/dx.doi.org/10.1111/j.1469-7610.2006.01618.x. [PubMed: 16492262]

Fallon T Jr, Schwab-Stone M. Determinants of reliability in psychiatric surveys of children aged 6–12. Journal of Child Psychology & Psychiatry & Allied Disciplines. 1994; 35:1391–1408. http://dx.doi.org/10.1111/j.1469-7610.1994.tb01282.x.

Ford T. Practitioner review: How can epidemiology help us plan and deliver effective child and adolescent mental health services? Journal of Child Psychology and Psychiatry. 2008; 49:900–914. http://dx.doi.org/10.1111/j.1469-7610.2008.01927.x. [PubMed: 18573144]

Gudmundsson E. Guidelines for translating and adapting psychological instruments. Nordic Psychology. 2009; 61:29–45. http://dx.doi.org/10.1027/1901-2276.61.2.29.

Hartman CA, Hox J, Mellenbergh GJ, Boyle MH, Offord DR, Racine Y, et al. Sergeant JA. *DSM-IV* internal construct validity: When a taxonomy meets data. Journal of Child Psychology and Psychiatry. 2001; 42:817–836. http://dx.doi.org/10.1111/1469-7610.00778. [PubMed: 11583254]

Ho, SMY.; Rochelle, TL.; Law, LSC.; Duan, W.; Bai, Y.; Shih, SM. Methodological issues in positive psychology research with diverse populations: Exploring strengths among Chinese adults. In: Pedrotti, JT.; Edwards, LM., editors. Perspectives on the intersection of multiculturalism and positive psychology Vol 7: Cross-cultural advancements in positive psychology. Dordrecht, The Netherlands: Springer; 2014. p. 45-57.http://dx.doi.org/10.1007/978-94-017-8654-6_4

Ivanova MY, Achenbach TM, Dumenci L, Rescorla LA, Almqvist F, Weintraub S, et al. Verhulst FC. Testing the 8-syndrome structure of the Child Behavior Checklist in 30 societies. Journal of Clinical Child and Adolescent Psychology. 2007; 36:405–417. http://dx.doi.org/10.1080/15374410701444363. [PubMed: 17658984]

Jensen PS, Goldman E, Offord D, Costello EJ, Friedman R, Huff B, et al. Roberts R. Overlooked and underserved: “action signs” for identifying children with unmet mental health needs. Pediatrics. 2011; 128:970–979. [PubMed: 22025589]

Kim ES, Yoon M. Testing measurement invariance: A comparison of multiple-group categorical CFA and IRT. Structural Equation Modeling. 2011; 18:212–228. http://dx.doi.org/10.1080/10705511.2011.557337.

Kovess V, Carta MG, Pez O, Bitfoi A, Koç C, Goelitz D, et al. Otten R. The School Children Mental Health in Europe (SCMHE) project: Design and first results. Clinical Practice & Epidemiology in Mental Health. 2015; 11(Suppl1: M7):113–123. http://dx.doi.org/10.2174/1745017901511010113. [PubMed: 25834631]

Kuijpers RCWM, Otten R, Krol NPCM, Vermulst AA, Engels RCME. The reliability and validity of the Dominic Interactive: A computerized child report Instrument for mental health problems. Child & Youth Care Forum. 2013; 42:35–52. http://dx.doi.org/10.1007/s10566-012-9185-7.

Kuijpers RCWM, Otten R, Vermulst AA, Engels RCME. Reliability and construct validity of a child self-report instrument: The Dominic Interactive. European Journal of Psychological Assessment. 2014; 30:40–47. http://dx.doi.org/10.1027/1015-5759/a000166.

Lahey BB, Applegate B, Waldman ID, Loft JD, Hankin BL, Rick J. The structure of child and adolescent psychopathology: Generating new hypotheses. Journal of Abnormal Psychology. 2004; 113:358–385. http://dx.doi.org/10.1037/0021-843X.113.3.358. [PubMed: 15311983]

Lahey BB, Rathouz PJ, Van Hulle C, Urbano RC, Krueger RF, Applegate B, et al. Waldman ID. Testing structural models of *DSM-IV* symptoms of common forms of child and adolescent psychopathology. Journal of Abnormal Child Psychology. 2008; 36:187–206. http://dx.doi.org/10.1007/s10802-007-9169-5. [PubMed: 17912624]

Marsh HW, Hau KT, Wen ZL. In search of golden rules: Comment on hypothesis testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu & Bentler (1999) findings. Structural Equation Modeling. 2004; 11:320–341. http://dx.doi.org/10.1207/s15328007sem1103_2.

McDonald, RP. Test theory: A unified treatment. Mahwah, NJ: Erlbaum; 1999.

Merikangas KR, Nakamura EF, Kessler RC. Epidemiology of mental disorders in children and adolescents. Dialogues in Clinical Neuroscience. 2009; 11:7–20. Retrieved from http://www.dialogues-cns.com/pdf/DialoguesClinNeurosci-11-7.pdf. [PubMed: 19432384]

Millsap RE, Kwok OM. Evaluating the impact of partial factorial invariance on selection in two populations. Psychological Methods. 2004; 9:93–115. http://dx.doi.org/10.1037/1082-989X.9.1.93. [PubMed: 15053721]

Muthén B. A general structural equation model with dichotomous, ordered categorical and continuous latent variable indicators. Psy-chometrika. 1984; 49:115–132. http://dx.doi.org/10.1007/BF02294210.

Muthén, LK.; Muthén, BO. Mplus user's guide. 4th. Los Angeles, CA: Author; 1998–2007.

O'Connell, ME.; Boat, T.; Warner, KE., editors. Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities Committee on Prevention of Mental Disorders and Substance Abuse Among Children, Youth and Young Adults: Research advances and promising interventions. Washington, DC: The National Academies Press; 2009. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK32775/

Reef J, Diamantopoulou S, van Meurs I, Verhulst F, van der Ende J. Child to adult continuities of psychopathology: A 24-year follow-up. Acta Psychiatrica Scandinavica. 2009; 120:230–238. http://dx.doi.org/10.1111/j.1600-0447.2009.01422.x. [PubMed: 19522882]

Revelle W, Zinbarg RE. Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. Psychometrika. 2009; 74:145–154. http://dx.doi.org/10.1007/s11336-008-9102-z.

Rutkowski L, Svetina D. Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. Educational and Psychological Measurement. 2014; 74:31–57. http://dx.doi.org/10.1177/0013164413498257.

Sass DA. Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. Journal of Psychoeducational Assessment. 2011; 29:347–363. http://dx.doi.org/10.1177/0734282911406661.

Schmitt N, Kuljanin G. Measurement invariance: Review of practice and implications. Human Resource Management Review. 2008; 18:210–222. http://dx.doi.org/10.1016/j.hrmr.2008.03.003.

Scott TJL, Short EJ, Singer LT, Russ SW, Minnes S. Psychometric properties of the Dominic Interactive assessment: A computerized self-report for children. Assessment. 2006; 13:16–26. http://dx.doi.org/10.1177/1073191105284843. [PubMed: 16443716]

Shojaei T, Wazana A, Pitrou I, Gilbert F, Bergeron L, Valla JP, Kovess-Masfety V. Psychometric properties of the Dominic Interactive in a large French sample. Canadian Journal of Psychiatry/Revue Canadienne de Psychiatrie. 2009; 54:767–776. [PubMed: 19961665]

Sijtsma K. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. Psychometrika. 2009; 74:107–120. http://dx.doi.org/10.1007/s11336-008-9101-0. [PubMed: 20037639]

Silverman WK, Ollendick TH. Evidence-based assessment of anxiety and its disorders in children and adolescents. Journal of Clinical Child and Adolescent Psychology. 2005; 34:380–411. http://dx.doi.org/10.1207/s15374424jccp3403_2. [PubMed: 16026211]

Sourander A, Haavisto A, Ronning JA, Multimäki P, Parkkola K, Santalahti P, et al. Almqvist F. Recognition of psychiatric disorders, and self-perceived problems. A follow-up study from age 8 to age 18. Journal of Child Psychology and Psychiatry. 2005; 46:1124–1134. http://dx.doi.org/10.1111/j.1469-7610.2005.00412.x. [PubMed: 16178937]

Steenkamp JEM, Baumgartner H. Assessing measurement invariance in cross-national consumer research. Journal of Consumer Research. 1998; 25:78–90. http://dx.doi.org/10.1086/209528.

Stone LL, Otten R, Ringlever L, Hiemstra M, Engels RCME, Vermulst AA, Janssens JMAM. The parent version of the Strengths and Difficulties Questionnaire: Omega as an alternative to alpha and a test for measurement invariance. European Journal of Psychological Assessment. 2013; 29:44–50. http://dx.doi.org/10.1027/1015-5759/a000119.

Valla, JP. Instruction manual for the Dominic Interactive. In: Valla, JP., editor. Dominic Interactive [CD-ROM]. Montreal, Canada: DIMAT; 2000.

Valla JP, Bergeron L, Bérubé H, Gaudet N, St-Georges M. A structured pictorial questionnaire to assess *DSM-III–R*-based diagnoses in children (6–11 years): Development, validity, and reliability. Journal of Abnormal Child Psychology. 1994; 22:403–423. http://dx.doi.org/10.1007/BF02168082. [PubMed: 7963075]

Valla JP, Bergeron L, Bidaut-Russell M, St-Georges M, Gaudet N. Reliability of the Dominic-R: A young child mental health questionnaire combining visual and auditory stimuli. Journal of Child Psychology & Psychiatry & Allied Disciplines. 1997; 38:717–724. http://dx.doi.org/10.1111/j.1469-7610.1997.tb01698.x.

Valla JP, Kovess V, Chan Chee C, Berthiaume C, Vantalon V, Piquet C, et al. Allès-Jardel M. A French study of the Dominic Interactive. Social Psychiatry and Psychiatric Epidemiology. 2002; 37:441–448. http://dx.doi.org/10.1007/s00127-002-0575-2. [PubMed: 12242621]

van Widenfelt BM, Treffers PDA, de Beurs E, Siebelink BM, Koudijs E. Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. Clinical Child and Family Psychology Review. 2005; 8:135–147. http://dx.doi.org/10.1007/s10567-005-4752-1. [PubMed: 15981582]

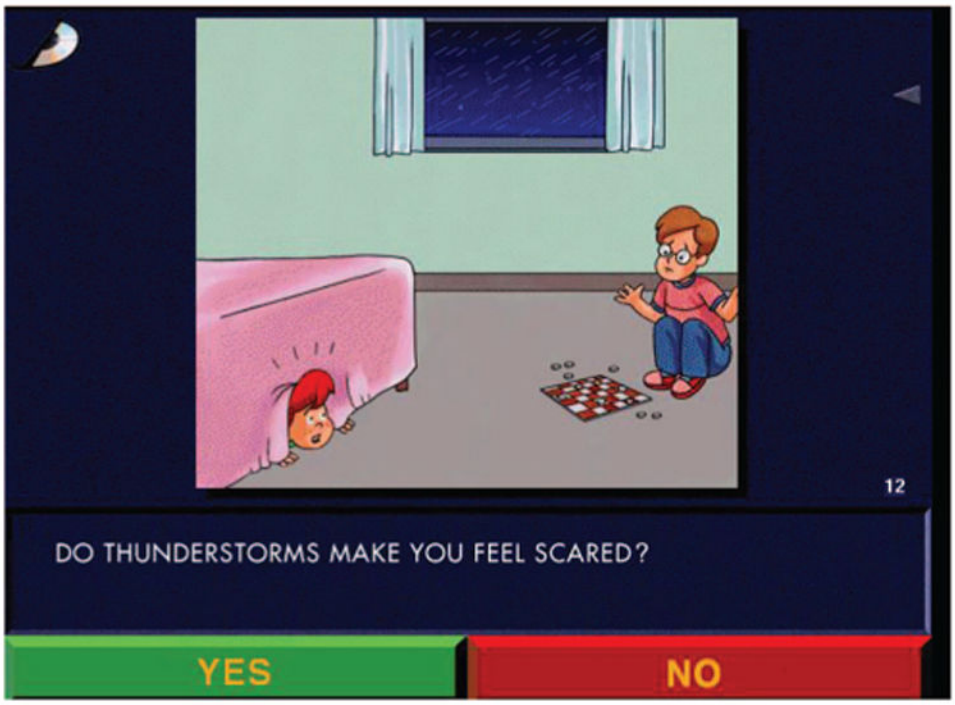


**Figure 1.**

Example of a symptom of the Dominic Interactive. Adapted from *Dominic Interactive* by J. P. Valla, 2000, Montreal: DIMAT. Copyright 2000 by J. P. Valla. See the online article for the color version of this figure.

**Table 1**

**Descriptive Statistics Separated by Country**

| | Country (*N* = 8,135) | | | | | | |
|---|---|---|---|---|---|---|---|
| **Variable** | **Bu** | **Ge** | **It** | **Li** | **Ne** | **Ru** | **Tu** |
| No. of children | | | | | | | |
| *n* | 1,385 | 894 | 757 | 1,278 | 1,503 | 1,397 | 921 |
| Age, years *M* (*SD*) | 8.8 (1.23) | 8.5 (1.24) | 8.2 (1.36) | 8.9 (1.19) | 9.0 (1.90) | 8.7 (1.23) | 8.7 (1.51) |
| Sex (%) | | | | | | | |
| Boys | 51.7 | 51.7 | 48.6 | 51.4 | 52.6 | 52.0 | 49.3 |
| Girls | 48.3 | 48.3 | 51.4 | 48.6 | 47.4 | 48.0 | 50.7 |
| Town size (%) | | | | | | | |
| *n* | 1,023 | 449 | 713 | 1,121 | 641 | 1,109 | 666 |
| Rural/village | 41.3 | 63.0 | 0.6 | 37.0 | 52.3 | 43.9 | 1.5 |
| Small/middle | 13.7 | 33.9 | 49.5 | 36.6 | 29.0 | 12.8 | 3.9 |
| Large | 45.0 | 3.1 | 49.5 | 26.4 | 18.7 | 43.3 | 94.6 |
| Nationality mother (%) | | | | | | | |
| *n* | 947 | — | 713 | 1,057 | 682 | 1,076 | 652 |
| Same country | 94.9 | — | 99.4 | 93.2 | 97.4 | 99.9 | 99.5 |
| EU country | 4.6 | — | 0.0 | 4.4 | 1.9 | 0.1 | 0.2 |
| Other | 0.4 | — | 0.6 | 2.4 | 0.7 | 0.0 | 0.3 |
| Nationality father (%) | | | | | | | |
| *n* | 840 | — | 703 | 943 | 625 | 967 | 650 |
| Same country | 95.2 | — | 98.2 | 92.6 | 95.8 | 98.6 | 99.5 |
| EU country | 3.9 | — | 0.4 | 4.2 | 2.4 | 0.4 | 0.3 |
| Other | 0.8 | — | 1.4 | 3.2 | 1.8 | 1.0 | 0.2 |
| Highest level of education parents | | | | | | | |
| *n* | 988 | 435 | 571 | 1,056 | 641 | 1,018 | 668 |
| Some primary or secondary | 7.5 | 5.9 | 0.2 | 6.0 | 0.9 | 12.6 | 40.2 |
| Secondary completed | 46.1 | 37.7 | 13.5 | 30.2 | 27.9 | 51.2 | 30.2 |
| College or technical school completed | 46.4 | 56.3 | 86.3 | 63.8 | 71.2 | 36.3 | 29.4 |

*Note*. Bu = Bulgaria; EU = European Union; Ge = Germany; It = Italy; Li = Lithuania; Ne = Netherlands; Ro = Romania; Tu = Turkey.

**Table 2**

**Reliability of the Scale Scores of the Dominic Interactive by Country**

| | Alpha (α) (*N* = 8,135) | | | | | | | Omega ($\omega_h$) (*N* = 8,135) | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Scale** | **Bu** | **Ge** | **It** | **Li** | **Ne** | **Ro** | **Tu** | **Bu** | **Ge** | **It** | **Li** | **Ne** | **Ro** | **Tu** |
| SPh | .69 | .57 | .54 | .60 | .57 | .64 | .66 | .86 | .80 | .75 | .77 | .80 | .80 | .84 |
| SAD | .65 | .72 | .63 | .59 | .75 | .67 | .59 | .80 | .86 | .79 | .76 | .87 | .83 | .75 |
| GAD | .77 | .80 | .70 | .71 | .75 | .77 | .74 | .87 | .89 | .83 | .83 | .87 | .89 | .85 |
| MDD | .84 | .83 | .76 | .82 | .78 | .86 | .83 | .92 | .93 | .88 | .91 | .91 | .94 | .91 |
| ODD | .79 | .73 | .75 | .72 | .73 | .79 | .71 | .91 | .87 | .88 | .88 | .88 | .92 | .89 |
| CD | .84 | .82 | .64 | .80 | .73 | .86 | .82 | .95 | .96 | .89 | .92 | .92 | .97 | .95 |
| ADHD | .85 | .84 | .83 | .86 | .85 | .87 | .86 | .93 | .93 | .92 | .94 | .93 | .95 | .94 |

*Note.* ADHD = attention-deficit/hyperactivity disorder; Bu = Bulgaria; CD = conduct disorder; GAD = generalized anxiety disorder; Ge = Germany; It = Italy; Li = Lithuania; MDD = major depressive disorder; Ne = Netherlands; ODD = oppositional defiant disorder; Ro = Romania; SAD = separation anxiety disorder; SPh = specific phobia; Tu = Turkey.

**Table 3**

**Correlations Between the Seven Factors of the Dominic Interactive (N = 8,135)**

| Scale | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| SPh | — | .34–.43 | .35–.45 | .27–.42 | .14–.35 | .05[a]–.30 | .18–.32 |
| SAD | .41 | — | .55–.66 | .39–.58 | .15–.37 | .08–.30 | .25–.39 |
| GAD | .41 | .63 | — | .48–.64 | .23–.41 | .11–.28 | .33–.47 |
| MDD | .36 | .49 | .58 | — | .37–.60 | .30–.52 | .53–.73 |
| ODD | .24 | .25 | .30 | .47 | — | .52–.64 | .54–.71 |
| CD | .19 | .19 | .17 | .38 | .57 | — | .46–.69 |
| ADHD | .23 | .30 | .34 | .59 | .63 | .57 | — |

*Note*. Total correlations are below the diagonal; range of correlations is above the diagonal. All correlations are significant at the $p < .01$ level. ADHD = attention-deficit/hyperactivity disorder; CD = conduct disorder; GAD = generalized anxiety disorder; MDD = major depressive disorder; ODD = oppositional defiant disorder; SAD = separation anxiety disorder; SPh = specific phobia.

[a]Significant at $p < .05$.

**Table 4**

**Goodness-of-Fit Indices of the Seven-Factor Structure of the Dominic Interactive (N = 8,135)**

| Model | Factor loadings and thresholds | $\chi^2$ | *df* | *p* | CFI | RMSEA | ΔCFI | ΔRMSEA |
|---|---|---|---|---|---|---|---|---|
| SPh | Free to vary | 382.08 | 182 | .000 | .961 | .031 | | |
| | Equal | 494.84 | 230 | .000 | .949 | .031 | −.012 | .001 |
| SAD | Free to vary | 559.52 | 133 | .000 | .957 | .053 | | |
| | Equal | 1638.35 | 175 | .000 | .853 | .085 | **−.104** | **.035** |
| GAD | Free to vary | 779.22 | 189 | .000 | .932 | .052 | | |
| | Equal | 1710.98 | 237 | .000 | .831 | .073 | **−.101** | **.021** |
| MDD | Free to vary | 923.85 | 532 | .000 | .980 | .025 | | |
| | Equal | 1759.50 | 610 | .000 | .942 | .040 | −.038 | .015 |
| ODD | Free to vary | 741.94 | 189 | .000 | .965 | .050 | | |
| | Equal | 1080.42 | 237 | .000 | .947 | .055 | −.017 | .005 |
| CD | Free to vary | 615.28 | 455 | .000 | .990 | .017 | | |
| | Equal | 773.21 | 527 | .000 | .984 | .020 | −.006 | .003 |
| ADHD | Free to vary | 2099.32 | 1064 | .000 | .967 | .029 | | |
| | Equal | 2855.68 | 1172 | .000 | .947 | .035 | −.020 | .006 |

*Note*. Boldface indicates noninvariant scales. ADHD = attention-deficit/hyperactivity disorder; CD = conduct disorder; CFI = comparative fit index; GAD = generalized anxiety disorder; MDD = major depressive disorder; ODD = oppositional defiant disorder; RMSEA = root mean square error of approximation; SAD = separation anxiety disorder; SPh = specific phobia.

*Psychol Assess*. Author manuscript; available in PMC 2017 May 01.

**Table 5**

**Factor Loadings and Thresholds of Items of Separation and Generalized Anxiety Disorder of the Dominic Interactive by Country**

| | | Factor loadings | | | | | | | Thresholds | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Disorder** | **Bu** | **Ge** | **It** | **Li** | **Ne** | **Ro** | **Tu** | **Bu** | **Ge** | **It** | **Li** | **Ne** | **Ro** | **Tu** | **ΔCFI** | **ΔRMSEA** |
| Separation anxiety disorder | | | | | | | | | | | | | | | | | |
| **3.** | **Worry about parents in car accident** | .31 | .48 | .39 | .26 | .51 | .33 | .34 | −.51 | −.10 | .61 | −.82 | .28 | −.67 | −.72 | **−.025** | .006 |
| **17.** | **Following parents everywhere** | .59 | .53 | .60 | .52 | .65 | .59 | .51 | .44 | 1.04 | .04 | 1.15 | 1.11 | .73 | .64 | **−.013** | .003 |
| 20. | 20. Refuse to go to school to be with parents | .45 | .50 | .40 | .37 | .60 | .49 | .39 | .86 | 1.14 | .79 | .93 | .58 | 1.34 | 1.13 | −.009 | .001 |
| 24. | Not able to sleep away from parents | .65 | .62 | .55 | .61 | .67 | .64 | .55 | .44 | .58 | .90 | .85 | .62 | .61 | .74 | −.005 | .002 |
| 27. | Nightmares about separation | .56 | .80 | .39 | .48 | .74 | .62 | .44 | .64 | .83 | .78 | .66 | .74 | .49 | .49 | −.008 | .001 |
| **30.** | **Worrying about separation** | .47 | .76 | .69 | .42 | .65 | .60 | .50 | −.51 | .28 | −.29 | −.72 | .30 | −.21 | −.63 | **−.023** | .004 |
| 34. | Feeling miserable away from parents | .65 | .69 | .77 | .69 | .80 | .73 | .68 | −.09 | .28 | .10 | −.04 | .75 | −.15 | −.24 | −.008 | .001 |
| 36. | Feeling sick when parents leave | .74 | .80 | .67 | .73 | .78 | .73 | .70 | .21 | .72 | .54 | .06 | .36 | −.08 | .20 | −.008 | .001 |
| Generalized anxiety disorder | | | | | | | | | | | | | | | | | |
| 5. | Finding it difficult to relax | .50 | .29 | .34 | .31 | .31 | .34 | .41 | .50 | .44 | .78 | .69 | .65 | .90 | .57 | −.001 | −.001 |
| 8. | Worrying about being bad at sports | .59 | .61 | .58 | .51 | .67 | .63 | .56 | .48 | .62 | .55 | .33 | .93 | .51 | .37 | −.004 | .000 |
| **10.** | **Nightmares** | .51 | .49 | .50 | .32 | .59 | .50 | .46 | .17 | .35 | .47 | .43 | .30 | .23 | .25 | **−.017** | .003 |
| **14.** | **Worrying about not having friends** | .47 | .64 | .38 | .62 | .66 | .68 | .65 | −.62 | −.16 | −.39 | −.50 | .70 | .03 | −.29 | **−.018** | .003 |
| **16.** | **Worrying about one's look** | .63 | .66 | .47 | .39 | .65 | .59 | .58 | .30 | .53 | .86 | −.39 | .53 | .46 | .24 | **−.025** | .004 |
| 19. | Worrying about having a car accident | .71 | .72 | .47 | .55 | .60 | .68 | .60 | −.30 | .10 | .24 | −.72 | .18 | −.24 | −.53 | −.005 | .000 |
| **21.** | **Worrying about getting lost** | .65 | .74 | .59 | .65 | .65 | .69 | .50 | .28 | .48 | −.28 | −.31 | .60 | −.01 | −.39 | **−.016** | .002 |
| **26.** | **Worrying about school grades** | .55 | .66 | .60 | .58 | .69 | .69 | .60 | −.51 | .00 | −.43 | −.53 | .77 | −.43 | −.68 | **−.017** | .003 |
| **32.** | **Feeling sick** | .51 | .58 | .42 | .29 | .53 | .40 | .51 | .49 | .62 | .80 | .52 | .72 | .46 | .52 | **−.011** | .001 |

*Note*. Boldface items are noninvariant. Bu = Bulgaria; Ge = Germany; It = Italy; Li = Lithuania; Ne = Netherlands; Ro = Romania; Tu = Turkey.

*Psychol Assess*. Author manuscript; available in PMC 2017 May 01.